Project Definition Document

End-To-End Solution using Globus Online Integrated with glideinWMS

1. Table of Contents

2. Approvals

glideinWMS Project Representative:	Signature:		Date:
	Print Name:	Burt Holzman	
	Title:		
Intensity Frontier Representative:	Signature:		Date:
	Print Name:	Lee Lueking	
	Title:		
REX Operations Representative:	Signature:		Date:
	Print Name:	Margaret Votava	
	Title:		
FermiGrid/Fermi Cloud Representative:	Signature:		Date:
	Print Name:	Steve Timm	
	Title:		
Computing Division Representative:	Signature:		Date:
	Print Name:	Gabriele Garzoglio	
	Title:		
Project Leader:	Signature:		Date:
	Print Name:	Parag Mhashilkar	
	Title:	Application Developer and System Analyst	

3. Document Change Log

Version	Date	Change Description	Prepared By
V 1.0	08/10/2011	First Version of the Document	Parag Mhashilkar

4. Project Proposal Lead

Project Leader : Parag Mhashilkar
Department : Computing Division
Group : SCF/GRID/DOCS

5. CD Strategy Document and Tactical Plans

Projects referred in this document are covered in following documents in the Fermilab's docdb -

- CEDPS is covered in tactical plan document : CD-doc-3826
- Fermi Cloud is covered in tactical plan document : CD-doc-4401
- glideinWMS is covered in tactical plan document : CD-doc-4394

6. Problem Statement

Different experiments as a part of the Intensity Frontier (IF) are ramping up their use of Grid resources. IF community is using glideinWMS to run their analysis workflows on the Grid and local clusters and has developed custom solutions to meet their data transfer needs.

As a part of this project, we intend to provide an end-to-end solution to the IF community that would integrate well with glideinWMS and their custom data transfer solutions leveraging on the modern techniques and solutions like Globus Online as provided by the Grid community.

7. Requirements

This section lists a set of functional and performance requirements that a proposed solution should satisfy.

1. The solution should get the user group membership information from the user credentials. The group is encoded as the name of the Fermilab VO subgroup for the given intensity frontier experiment. The corresponding local group ID is obtained from IF NIS servers.
2. The solution should preserve the file ownership properties of a transferred file. This includes username and the user group of the owner transferring the file. In case the user transferring the file belongs to multiple groups, membership should be determined based on the user credential that initiated the transfer.
3. The solution should be easy to integrate with glideinWMS.
4. The solution should support strong security like Kerberos or X509
5. The solution should utilize IF project disk areas mounted via Bluearc as a destination storage area.

6. The solution should either work for the different IF experiments or it should be convenient enough to replicate for different IF experiments.
7. It should be easy to add and/or remove support to individual IF experiments.
8. The solution should be scalable enough to the increasing demand.
9. Fermigrid members should have administrative rights on the machine.
10. Expected data transfer rates are not known at the time of writing this document.

8. Project Description and Goals

This project is an investigative effort that focuses on providing the IF community a working prototype of an end-to-end data transfers solution using Globus Online with glideinWMS. Globus Online is a service that optimizes data transfer parameters and provides convenient web-based monitoring interfaces. The work done is Fermilab's contribution as part of multi-institution collaborative project, Center for Enabling Distributed Petascale Science (CEDPS).

Setup a test bed for testing gridftp service
One of the goals for this project is to setup a test bed for testing gridftp service by the IF and REX community that will preserve the ownership information of the user transferring the files. To achieve this, investigate Fermi Cloud and use the Fermi Cloud services to create a Virtual Machine for deploying the test bed.

Develop tools to generate authentication/authorization list for accessing gridftp service
IF VO membership information is available in the Fermilab VOMS service. List of users affiliated with individual IF experiments, their UID and GID are available through the NIS server maintained by FES group. One of the goals for this project is to utilize and integrate this information to develop solution and tools creating the authentication/authorization list for accessing the gridftp service by different IF experiments.

The project understands that the number of individual IF experiments can change. Also, to simplify the operations, different VMs that would host the gridftp service should be identical.

Register the gridftp services with Globus Online Service
In order to use Globus Online for transferring the files, the gridftp service deployed on VMs should be correctly registered with the Globus Online. The project will appropriately register the gridftp servers requested today and provide documentation for the registration of future possible instances of gridftp.

Develop condor-globusonline-transfer-plugin
Newer version of Condor has plugin architecture that lets the users write scripts to support custom file transfer protocols. Since condor is used as WMS inside glideinWMS, one of the goals is to write a plugin that lets condor handle the Globus Online transfers natively.

Integrate condor-globusonline-transfer-plugin with glideinWMS
In order for the condor in glideinWMS to work with custom transfer plugins, glideinWMS should be enhanced to support condor configuration with custom file transfer plugins.

Deploy glideinWMS for end-to-end testing

To enable end-to-end testing and the commissioning of a functional system for the IF community, one of the goals for this project is to deploy glideinWMS with support for custom condor file transfer plugins.

9. Project Scope

This section describes what is in scope and what is out of scope for the project.

9.1 What is in Scope

Scope of the project includes –

- Utilize Fermi Cloud resources to deploy services
- Develop administrative tools to provide authentication/authorization to the deployed services
- Develop tools to interface the solution with glideinWMS and with Globus Online
- Provide customer support to the stakeholders until the provided solution is transitioned to operations
- Transition the tools developed as part of this project to REX operations
- Transition the operations of Fermi Cloud Virtual Machines and images to the FermiGrid

9.2 What is out of Scope

Tasks/items not listed in the section 9.1 are out of scope of this project.

10.Project Deliverables and Milestones

Some of the high level milestones and deliverables for the project are –

Milestones/Deliverables	Requester/ Stakeholder	Planned For
Setup a test bed for testing gridftp service	IF, REX	06/29/2011
Develop tools to generate authentication and authorization list for accessing gridftp service	IF, Fermi Cloud	06/27/2011
Create gridftp service endpoints for each IF experiment	IF, Fermi Cloud	08/28/2011
Register the gridftp services with Globus Online Service	IF	08/29/2011
Develop condor-globusonline-transfer-plugin	CEDPS, glideinWMS	08/03/2011
Integrate condor-globusonline-transfer-plugin with glideinWMS	CEDPS, glideinWMS	08/19/2011
Deploy glideinWMS for end-to-end testing and commissioning	IF, REX, CEDPS	08/26/2011
Close the Project	CEDPS	09/30/2011

11.Project Organizational Structure

The program of work and efforts for this project are partially sponsored by CEDPS activity in the Fermilab.

11.1 Sponsor(s)

Effort for this project is partially funded by the CEDPS activity in the Fermilab.

11.2 Stakeholders

glideinWMS Project : Burt Holzman
IF VO : Lee Lueking
REX Operations : Margaret Votava
FermiGrid/FermiCloud : Steve Timm
Computing Division : Gabriele Garzoglio

11.3 Project Team

Name	Project Role	Effort (FTE Equivalent)
Burt Holzman	Project Manager	10% (07/01/2011 – 09/30/2011)
Parag Mhashilkar	Project Leader	30% (07/01/2011 – 09/30/2011)
Marko Slyz	Team Member	5% (07/01/2011 – 09/30/2011)
Xi Duan	Summer Student	50% (07/15/2011 – 09/15/2011)

12.Preliminary Project Plan / Statement of Work

12.1 WBS

- Project Documentation
  - Communicate with the stakeholders
  - Gather requirements
  - Write project definition document
- Develop tools to generate authentication and authorization list for accessing gridftp service
  - Script to generate gridmapfile for the gridftp service
  - Script to generate local users & groups by contacting the IF NIS
- Setup a test bed for testing gridftp service
  - Create a test VM on Fermi Cloud if-gridftp.fnal.gov
  - Deploy gridftp service on if-gridftp.fnal.gov
  - Deliverable: if-gridftp.fnal.gov VM with gridftp service that correctly preserves the group membership of transferred files
- Deploy one gridftp end point per IF experiment
  - Create a VM with the gridftp service deployed for one experiment
  - Clone the VM
  - Deploy multiple clones, one per IF experiments
- Register the gridftp services with Globus Online
  - Register the gridftp servers used by IF experiments as individual GO end points
- Develop condor-globusonline-transfer-plugin
  - Develop condor-globusonline-transfer-plugin
  - Test condor-globusonline-transfer-plugin
  - Release condor-globusonline-transfer-plugin
- Integrate condor-globusonline-transfer-plugin with glideinWMS

    - Integrate condor-globusonline-transfer-plugin with glideinWMS
    - Test the integration work
    - Release glideinWMS with support for globusonline-transfer-plugin
- Deploy glideinWMS for end-to-end testing
    - Deploy glideinWMS
    - Test the deployment
    - Commission the deployment
- End-To-End Test for running jobs on IF glideinWMS deployment
    - Do end-to-end testing using glideinWMS and Globus Online
- Close the Project
    - Write the closing document
    - Deliverable: Closing document

12.2 Computer Security Considerations

Hardware and software/services required for the project will be hosted at Fermilab and will follow the Computer security policies of Fermilab. The Globus Online Service is jointly hosted and maintained by the Globus team at Argonne National Laboratory and University of Chicago. Gridftp service that acts as the globusonline.org endpoint is deployed in the virtual machine hosted by the Fermi Cloud.

12.3 Operations Responsibilities at Close of Project

Hardware, virtual machine and the operating system that hosts gridftp endpoint will be hosted at Fermilab and will follow the Computer security policies imposed by Fermilab. Hardware and operating system for the virtual machine is supported by the FermiGrid operations group and will continue to remain the same. VDT installation required for the deployment of the gridftp and gridftp operations will be maintained and supported by the REX group.

13.Estimated Resource Requirements

This section describes the personnel and equipment cost for the project.

13.1 Personnel Cost

Personal cost is listed in the Project Organization structure section above.

13.2 Hardware Cost

Estimating the hardware cost is outside the scope of this document. The hardware and the resources utilized for this project are hosted by the Fermi Cloud project. Costing for the resources will be done by the Fermi Cloud project and presented to the IF community separately.

14.Supporting Documentation

The project web home page is https://cdcvs.fnal.gov/redmine/projects/cedps-glideinwms

This link includes documentation for two individual subprojects, IF Gridftp setup and Condor-globusonline-transfer-plugin.

15.Project Risks, Issues, and Assumptions

Risk	Impact Level	Risk Plan Actions
Support for Globus Online being dropped	Medium	Globus Online Services are maintained by the Globus Group at Argonne National Laboratory & University of Chicago. In case the Globus Online Services project is discontinued, the proposed solution will still work since it depends on the availability of the gridftp service. A new condor file transfer plugin needs to be written or the existing working model needs to be adjusted to use gridftp end point natively.
Delay in migrating the ownership of condor-globusonline-transfer-plugin the to Condor community	Medium	This will need working with the glideinWMS team so that they continue supporting the plugin as part of the glideinWMS project.
Delay in closing the project by 09/30/2011	High	This project is funded by the CEDPS activity in the Fermilab. If the work on this project is delayed, an alternate source of funds to support this activity past its original completion date of 09/30/2011 will be required.